Postal Regulatory Commission
Submitted 9/15/2017 3:32:22 PM
Filing ID: 101649
Accepted 9/15/2017

BEFORE THE
POSTAL REGULATORY COMMISSION

Inquiry Concerning City Carrier Costs	:	**Docket No. PI2017-1**

COMMENTS OF UNITED PARCEL SERVICE, INC. ON NOTICE AND ORDER ESTABLISHING DOCKET CONCERNING CITY CARRIER SPECIAL PURPOSE AND LETTER ROUTE COSTS AND TO SEEK PUBLIC COMMENT
(September 15, 2017)

United Parcel Service ("UPS") respectfully submits these comments in response to Commission Order No. 3926 (May 31, 2017) seeking comments to evaluate the Postal Service's progress in its ongoing efforts to update its city carrier cost models and data collection capabilities in accordance with Order No. 2792.

The issues addressed in this docket are critically important. They concern the failure of the Postal Service's costing models to identify the full scope of city carrier delivery costs that are attributable to competitive products. City carrier costs represent the largest body of costs in Postal Service operations, and it is city carriers who have for years been asked to deliver more and more competitive products.[1] The Postal Service's costing models for this activity, however, reflect assumptions from a different era, when regular mail delivery overwhelmingly drove the Postal Service's operations. Today, as regular mail delivery has dramatically declined, the Postal Service is increasingly focused on delivering competitive products.

[1] Together, Cost Segments 6 & 7 had total costs of more than $15.9 billion in FY16, more than any other cost segment.

Indeed, in the years since the Postal Accountability and Enhancement Act of 2006 was enacted, the Postal Service's parcel delivery business has grown from relatively modest volumes to the point that the Postal Service today delivers more packages to American homes than any other enterprise. The assumptions and special studies built into the Postal Service's legacy costing models fail to account for this dramatic change in its business.

Over the last seven years, market dominant volumes have shrunk by 15%, and city carrier delivery costs attributed to market dominant products have declined by roughly the same amount. As the below table shows, during the same period competitive product volumes have nearly *quadrupled*, while city carrier costs attributed to them has not kept pace.

Table 1: Volume and City Carrier Delivery Costs, FY09 – FY16

	Domestic Market Dominant Mail		Domestic Competitive Products	
Fiscal Year	Domestic Market Dominant Mail Volume (billions of pieces)	City Carrier Delivery Costs Attributed to Domestic Market Dominant Mail (millions)	Domestic Competitive Product Volume (billions of pieces)	City Carrier Delivery Costs Attributed to Domestic Competitive Products (millions)
2009	175.7	7,498	1.1	300
2016	149.8	6,384	4.3	825
% Growth ('09 to '16)	-15%	-15%	296%	175%

Sources: Cost Segment and Components Reports, Revenue Pieces and Weight Reports for FY09 and FY16.

This tendency for attributable costs to lag growth in competitive product volumes is especially dubious in light of mounting evidence that the Postal Service is increasingly relying on overflow and parcel-only routes to accommodate the growing volume of these products. Because such routes are caused by parcel volumes, they logically should be attributed entirely (or nearly entirely) to competitive products.

Proposal Thirteen updated the cost model that had been used to attribute city carrier delivery costs since 2002. The 2002 study was a one-time effort that reflected those conditions prevailing at the time. Reliance on this study thus effectively locked in place costing procedures that became increasingly out of touch with ongoing operations. While Proposal Thirteen updated the study, it replicated the Postal Service's methodological error and thus "freeze[s] in place an outdated approach to cost accounting" for the years to come.[2] Proposal Thirteen relied on data collected over two special studies conducted in 2013 and 2014, both of which lasted just *two weeks* in length and covered only 300 ZIP codes (out of more than 10,000). Most important, neither study was conducted during the peak season when competitive product volume surges, straining the Postal Service's delivery systems.[3]

Three years later, the existing cost-attribution models are still based on data gathered from 2014, even as Postal Service parcel volumes continue to grow rapidly. Despite partial fixes, like the still-pending Proposal Four, the underlying issue remains unaddressed—key inputs to the cost model are still based on data collected in the past, and there is no way to test or monitor how the Postal Service's increasingly diverse mail stream (*e.g.*, new products) is affecting delivery costs.

As the Commission is aware, UPS has proposed a modern approach to cost attribution that is capable of reflecting the ongoing changes in the Postal Service's

---

[2] United Parcel Service Comments on Postal Service Proposal Thirteen Regarding City Carrier Street Time Costs, Dkt. No. RM2015-7 (Mar. 18, 2015), at 2.

[3] *See* Petition of the United States Postal Service for the Initiation of a Proceeding to Consider Proposed Change in Analytical Principles (Proposal Thirteen), Dkt. No. RM2015-7 (Dec. 11, 2014), at 1-4. Proposal Thirteen also reflected two critical assumptions about cost causation that are at best untested, and at worst highly questionable – namely, that letter mail volumes have no effect on parcel delivery times, and that rapidly growing parcel volumes have no effect on letter mail delivery times.

business, including the increased cost burden imposed by parcel delivery. Specifically, UPS proposed using a "single equation" model that would rely on operational data collected in the ordinary course of business.[4] This modeling approach has the potential to improve estimates of variability while reducing the need for expensive special data collection studies.[5] It also approaches cost causation in a more agnostic, data driven way, making it able to capture accurately the complex ways in which rapid growth in competitive product volumes are reshaping the delivery environment.

The Postal Service criticized UPS's proposed model, arguing that it suffered from data quality problems.[6] The Postal Service made this criticism notwithstanding the fact that its own approach had even more serious limitations.

Despite these criticisms, however, the Commission urged the Postal Service to investigate the feasibility of a single equation city carrier letter route cost model.[7] That investigation has been underway for two years. Although the Postal Service has made progress by incorporating new and more reliable sources of parcel volume data, there is still no clear path toward operationalizing this new approach. Instead of embracing the approach, the Postal Service has resisted the use of a "single equation" model, claiming that limitations in the quality of its own data prevent it from being implemented. Rather than using these limitations as a reason not to move forward with a superior approach,

---

[4] United Parcel Service Comments on Postal Service Proposal Thirteen Regarding City Carrier Street Time Costs, Dkt. No. RM2015-7 (Mar. 18, 2015), at 23-28.

[5] *See* United Parcel Service, Inc.'s Comments and Supplemental Report Regarding Proposal Thirteen, Dkt. No. RM2015-7 (June 8, 2015), at 2-4.

[6] Reply Comments of the United States Postal Service in Response to March 18th Comments, Dkt. No. RM2015-7 (May 13, 2015), at 23-41.

[7] Order Approving Analytical Principles Used in Periodic Reporting (Proposal Thirteen), Dkt. No. RM2015-7 (Oct. 29, 2015), at 64-66.

the Postal Service should be determined to solve them. Otherwise, no one can be confident that costing models are properly attributing costs. This likely means, for example, that market dominant mailers are forced to bear a disproportionate share of delivery costs.

Indeed, preliminary results from the Postal Service's own prototype single-equation models indicate that costs are being systematically under-attributed to competitive products, with under-attribution (relative to Proposal Thirteen) ranging from 29% to 57% (or anywhere from $242 to $471 million). In other words, the results presented by the Postal Service appear to confirm that market dominant products are indeed responsible for a disproportionate share of costs. This result is consistent with UPS's concern, expressed in 2015, that Proposal Thirteen would "systematically understate attribution of costs to parcels for the indefinite future, even as the Postal Service devotes increasingly more resources (and therefore costs) to that business."[8] As discussed below, concerns also persist about the Postal Service's treatment of special purpose routes, which is another subject addressed by this docket.

The data quality concerns raised by the Postal Service as a reason not to move forward are not fatal to the use of a single equation model. As detailed below, there are potential solutions to *every* concern raised by the Postal Service. The Commission should continue proceedings in this docket until those problems are addressed and a single equation model is implemented. UPS would welcome a technical conference on the outstanding issues and the proposed solutions detailed below.

[8] United Parcel Service Comments on Postal Service Proposal Thirteen Regarding City Carrier Street Time Costs, Dkt. No. RM2015-7 (Mar. 18, 2015), at 3.

I. COMMENTS REGARDING A POTENTIAL SINGLE EQUATION CITY CARRIER LETTER ROUTE COST MODEL

In its August 18th filing in this docket, the Postal Service reports the results of eight prototype single-equation models that varied in the time period covered by the data, in their treatment of parcel volumes, and in the treatment of categories of ZIP codes.[9] The Postal Service did not present potential cost impacts for any of its models, but it is possible to estimate those impacts by inputting the estimated variabilities from the prototype models into the current Cost Segment 6 & 7 models.[10]

These admittedly preliminary results imply that current costing methods systematically under-attribute costs to competitive products by amounts ranging from 29% to 57%, as shown in Table 2. For example, the first row in Table 2, corresponding to the first prototype model presented in the Postal Service's report, implies that competitive product attributable costs in FY16 exceeded those implied by the current costing methodology by $403 million.

[9] *See generally* Report on Research Into the Ability of a Top-Down Model To Accurately Estimate City Carrier Street Time Variabilities, Dkt. No. PI2017-1 (Aug. 18, 2017) ("USPS Report").

[10] This also requires an adjustment to the cost pool formation step of the process, such that there is a single street time cost pool.

Table 2: Cost Segment 6 & 7 Cost Comparison: Established Model, Proposal Four, and 8 Prototype Top-Down Models

Prototype Top-Down Model			Total Domestic Competitive Mail and Services Attributable Costs			% Change in Competitive Attributable Costs	
Date Used	Treatment of Parcels	Separate Equation for FSS zones?	Protoypical Model (Illustrative)	Established Model (As Filed)	Proposal Four (RM2017-8)	Relative to Established Model	Relative to Proposal Four (As Filed)
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
July	Split	No	1,296	825	971	57%	33%
July	Split	Yes	1,241	825	971	50%	28%
July	Unified	No	1,089	825	971	32%	12%
July	Unified	Yes	1,066	825	971	29%	10%
July & September	Split	No	1,263	825	971	53%	30%
July & September	Split	Yes	1,239	825	971	50%	28%
July & September	Unified	No	1,119	825	971	36%	15%
July & September	Unified	Yes	1,129	825	971	37%	16%

Sources: Petition for Proposal 4; CHIR No. 1 Response (Docket PI2017-1); USPS-PI2017-1/2; USPS-FY16-32
Notes: Reported costs reflect the estimated total for City Carrier Office and Street costs (C/S 6&7).
Each row corresponds to one of the prototype top-down models discussed in the August 2017 USPS Report, with the details specified in columns [1] through [3].
[7]: ([4] - [5]) / [5]
[8]: ([4] - [6]) / [6]

While the prototype models are by definition preliminary, it is striking that all eight approaches attribute far greater amounts of cost to competitive products than the current established methodology. These results confirm that it is urgent for the Postal Service to take steps to implement a proper "single equation" cost model for city carrier street time.[11] The current model used by the Postal Service simply does not work.

The Postal Service has raised some potential concerns with implementing a single-equation approach for city carrier street time. As a preliminary matter, however, it is critical to recognize that the single equation models tested by the Postal Service should not be measured against a perfect model. Rather, they should be measured against *Proposal Thirteen*, which undeniably suffers from significant limitations. Moreover, as explained below, there are ways to address each of the concerns raised

[11] Proposal Four (Dkt. No. RM2017-8) is certainly a step in the right direction towards capturing the increasingly important role parcels play in the delivery system. But it is ultimately a band-aid that does not solve the fundamental underlying issues with Proposal Thirteen.

by the Postal Service. The Commission should direct the Postal Service to consider each of these improvements and to make its best efforts to upgrade to a "single equation" cost model in the near term. The Commission initially requested the Postal Service to submit a plan to vet the model in 2015, two years ago. Comparing the FY14 data (which was the most recent data available at the time of the RM2015-7 order) with the FY16 data (the most recent annual data), competitive product volume has grown 34.9% while market dominant mail volumes have decreased by 1.4%.[12] Continued delay will only result in additional costing distortions that force market dominant mailers to pay for parcel delivery at an accelerating pace.

A. Multicollinearity

The Postal Service reports a high degree of multicollinearity in the prototype single-equation models it has estimated thus far.[13] It opines that traditional remedies for multicollinearity—notably adding additional observations—"do not appear to provide a solution for multicollinearity."[14]

Multicollinearity can occur when a model is used in an attempt to estimate a large number of coefficients from a dataset that does not have sufficient variation to estimate those coefficients precisely. But multicollinearity is ultimately a data issue, and not a modeling issue.[15] In other words, purported concerns about multicollinearity provide no

---

[12] *Compare* UNITED STATES POSTAL SERVICE, REVENUE, PIECES, AND WEIGHTS BY CLASSES OF MAIL AND SPECIAL SERVICES FOR FISCAL YEAR (2014), at 3-5, *with* UNITED STATES POSTAL SERVICE, REVENUE, PIECES, AND WEIGHTS BY CLASSES OF MAIL AND SPECIAL SERVICES FOR FISCAL YEAR (2016), at 3-5.

[13] *See* USPS Report at 38-39 ("A number of different indicators suggest a high degree of multicollinearity exists, undermining the equation's ability to produce reliable estimates of the effects of different types of volume on street time.").

[14] *Id*. at 14-15.

[15] *See* PAUL ALLISON, MULTIPLE REGRESSION: A PRIMER 149 (1999).

reason to question the value of using a single equation model. Rather, they at most reflect the limitations of the data that the Postal Service chose to use in its first attempts using that modeling approach. And in any event, based on the regressions the Postal Service has presented, the prototype model results appear to be sufficiently precise for attributing costs notwithstanding multicollinearity concerns.[16]

The simplest way to address multicollinearity is often to increase the amount of data—or more precisely, the amount of variation—used to estimate the model. An analyst can often address multicollinearity by using a dataset that contains more observations.[17] In this case, the Postal Service states that it considered the possibility of adding more data, but concluded additional data would "not appear to provide a solution for multicollinearity."[18] The Postal Service, however, appears to have only considered expanding its dataset from a single month (July 2016) to two months (adding September 2016), while continuing to limit the geographic coverage of the dataset to 300 ZIP codes.[19] That is not the only option.

Indeed, a key advantage of an approach that uses operational data is the availability of data covering more than 10,000 ZIP codes.[20] Rather than concluding, based on a single expansion of its dataset, that "the addition of data will likely have only a minimal impact on multicollinearity,"[21] the Postal Service should consider whether

---

16 *See* Appendix A.

17 *Id*.

18 USPS Report at 40.

19 *Id*. at 32-37.

20 *Id*. at 17.

21 *Id*. at 14-15.

exploiting a dataset with much broader geographic coverage—which is readily available—would resolve multicollinearity concerns.

The Postal Service could also use data collected throughout the calendar year.[22] The Postal Service only utilized data from July or September, which fails to exploit the differences in volume trends that occur in other months. For example, parcel and first-class mail volumes are known to increase in December, at the same time that standard mail volumes are decreasing, relative to their October-November peak. Inclusion of data from these and other months is likely to provide additional variation that should improve the performance of the model.

UPS has analyzed the multicollinearity issue present in the prototype single-equation models.[23] While the Postal Service has tested the effects of adding September 2016 data to July 2016 data, UPS's analysis suggests there would be a more impactful reduction in multicollinearity by adding data from additional ZIP codes.[24] UPS's analysis also demonstrates that the incorporation of data from other months throughout the year would add significant variation to the dataset, which would improve the robustness of the model.[25] The Commission should instruct the Postal Service to consider these solutions to its multicollinearity concerns.

B. Collections Volume

The Postal Service has concluded that the Product Tracking and Reporting ("PTR") system "holds the potential to provide reliable daily volumes for both parcels

---

22 *Id*. at 32.

23 The results of this analysis are attached in Appendix A.

24 *See* Appendix A.

25 *Id*.

and accountables and, quite possibly, provide the split between in-receptacle and deviation parcels."[26] This conclusion leaves collections mail as the last piece of data required to implement a "single equation" model. UPS therefore believes a more thorough investigation into potential sources of collection mail data is warranted.

UPS has four proposals along these lines, discussed below. Consistent with the Commission's traditional position that the Postal Service should proactively explore ways to collect relevant data,[27] the Commission should order the Postal Service to evaluate these four approaches and provide a report on the feasibility of each. Should these approaches prove feasible, there is then the follow-up question on how often the Postal Service should implement them. UPS proposes the Postal Service implement whichever approach works best on an annual basis.[28]

***The "Linear Feet" Approach*.** UPS has proposed that the Postal Service could institute changes that allow it to "measure the weight or linear feet of collection mail by route at the processing center *before* it is pooled."[29] This solution is consistent with the Postal Service's description of how mail is collected from customer receptacles, as Postal Service carriers drop off collections mail at the closest Post Office before it is

---

[26] USPS Report at 38.

[27] *See, e.g.*, Order Approving Analytical Principles Used in Periodic Reporting (Proposal Thirteen), Dkt. No. RM2015-7 (Oct. 29, 2015), at 65-66.

[28] It is likely that a yearly update would not be cost prohibitive. The figures provided by the Postal Service indicate that the cost of gathering this information for one ZIP code for one day comes to approximately 54 dollars. See Response Of The United States Postal Service To Commission Order No. 2792, February 16, 2016 at 12. The figure of approximately 54 dollars is equal to 14 routes per ZIP code times 3 minutes per route times an hourly carrier wage of $41.32, plus a half hour of supervisor time at an hourly wage of $51.12.

[29] United Parcel Service, Inc.'s Reply Comments to Response of United States Postal Service to Commission Order No. 2792, Dkt. No. RM2015-7 (March 4, 2016), at 6-7 (emphasis added).

bundled up and sent to a processing center.[30] This practice suggests that there are opportunities to conduct the measurement before the mail is pooled.

After UPS proposed this solution in March of 2016, the Postal Service cited potential cost concerns, replying that UPS "provides no evidence or analysis of why it would be cheaper for the Postal Service to count linear feet of collection mail at mail processing centers rather than to do the same thing, as the Postal Service proposed, at delivery units."[31] These costs could be reduced, however, by taking measurements for only a *sample* of ZIP code days. This would mirror the approach to collections volume in Proposal Thirteen itself.[32] The field procedures employed in that study for gathering collection mail volume were much simpler than those employed for gathering parcels and accountables data, which suggests that the latter portion of the study accounted for a disproportionate share of the data collection costs incurred in that effort.[33]

The "Reprogramming" Approach. The Postal Service has suggested that it might be possible to reprogram the scanners used by city carriers to allow them to record collection volumes.[34] The Postal Service has not offered any specific proposal for this approach, but this approach may solve the issue as well.

The "Form 3999" Approach. UPS has also previously suggested that the gathering of collection volume data could be made part of the route evaluation process

---

[30] *See* Responses of the United States Postal Service to Questions 1-9 of Chairman's Information Request No. 3, Dkt. No. PI2017-1 (Sept. 5, 2017), at 8 (response to Question 5).

[31] *See* Response of the United States Postal Service to UPS Pleading Regarding Commission Order No. 2792, Dkt. No. RM2015-7 (March 11, 2016), at 8.

[32] *See* Report on the City Carrier Street Time Study, Dkt. No. RM2015-7 (Dec. 11, 2014), at 27.

[33] *See id*. at 27-40 (collection volume study), 91-101 (package and accountables study).

[34] USPS Report at 2.

whose results are recorded in the Form 3999 database.[35] The Postal Service has argued that such an expansion of the route evaluation process would be prohibitively expensive.[36] The development of a workable estimation dataset, however, would not require the annual gathering of collection volume data for all 140,000 city carrier routes. Again, a suitably diverse and representative sample of ZIP codes would suffice.

The "CCCS" Approach. The Postal Service could also piggyback the gathering of collection volume data onto the City Carrier Costing System (CCCS). This system collects data annually for a random sample of route-days.[37] Data collection personnel visit the facilities with which the sampled routes are associated in order to record the composition of the mail delivered on that route.[38] If they were to gather facility level collection mail data while they are there, the result would be a collection mail volume dataset covering approximately 1,000 ZIP code days per year. Such a dataset could be linked to volume and time data drawn from the Postal Service's operational systems, producing a rich dataset with sufficient geographical and temporal variation to support the development of a single-equation approach.[39]

---

[35] United Parcel Service Comments on Postal Service Proposal Thirteen Regarding City Carrier Street Time Costs, Dkt. No. RM2015-7 (Mar. 18, 2015), at 21.

[36] Reply Comments of the United States Postal Service in Response to March 18th Comments, Dkt. No. RM2015-7 (May 13, 2015), at 32.

[37] Responses of the United States Postal Service to Questions 1-10 of Chairman's Information Request No. 2, Dkt. No. PI2017-1 (Jul. 25, 2017), at 6-8.

[38] Responses of the United States Postal Service to Questions 1-7 of Chairman's Information Request No. 1, Dkt. No. PI2017-1 (Jun. 30, 2017), at 17.

[39] Using this or other sampling-based approaches will, all things equal, shrink the estimation dataset and could thus exacerbate any potential multicollinearity concerns. If the smaller dataset is based on a random sample of ZIP code-days, all things will *not* be equal. The extent to which one might be concerned about multicollinearity worsens in a smaller sample depends in large part on the geographic and temporal coverage provided by the dataset. As we explained above, the Postal Service's explorations thus far have not yet come close to exploiting the available variation.

C. Accountables Volume

The Postal Service has expressed concern that due to the relatively low accountables volumes and the presence of non-volume related factors influencing street hours, the top-down approach may not be able to estimate accountables variabilities.[40] It is true that accountables are a fairly infrequent mail type. A typical route may handle only one or two accountable pieces per day.[41] Given the very low volumes of accountables, problems such as those encountered by the Postal Service are to be expected, but they can also be remedied.

Inspection of the "Form 3999" dataset produced in RM2017-8 suggests that there are some delivery routes with much higher accountable time and likely volumes.[42] Expansion of the prototype model dataset to include such routes—either through use of a larger or richer sample of ZIP code days or an explicit plan to over-sample high accountables ZIP codes—could provide the additional variation needed to produce acceptably precise estimates of marginally accountable delivery times.

Another approach would be to take accountables out of the single-equation model altogether. In Proposal Four, for example, the Postal Service proposed a methodology for calculating the size of the deviation parcel and accountables delivery cost pool.[43] This same methodology could be adapted to form an "accountables only"

---

[40] USPS Report at 15.

[41] *Id*.

[42] The average accountable delivery time is 3.7 minutes and the 95th percentile is 13.8 minutes.

[43] Petition of the United States Postal Service for the Initiation of a Proceeding to Consider Proposed Changes in Analytical Principles (Proposal Four), Dkt. No. RM2017-8 (Jun. 30, 2017), at 2-3.

cost pool.[44] The time associated with the "accountables only" cost pool could then be excluded from the model for the remaining shapes of mail. Such an exclusion would be an approximation that is similar to approximations used under the currently accepted methodology for handling "allied" activities.[45]

The Commission should order the Postal Service to consider these potential improvements to the model's treatment of accountables. Accountables do not make up a significant portion of mail volume and should not inhibit implementation of a single-equation approach.

II. COMMENTS REGARDING THE COST MODEL FOR SPECIAL PURPOSE ROUTES

In Order No. 3926, the Commission also sought to "review" the SPR cost model for street time."[46] The Postal Service has acknowledged that "the relative proportions of the various activities performed by Special Purpose Routes may have changed over the years."[47] And indeed there has been a pronounced change in the makeup of SPRs between 2010 and 2016, a development which raises questions about the ongoing validity of the legacy costing methods still being applied to SPRs.

---

[44] The special study on which the current delivery cost model is based recorded delivery times separately for deviation parcels, in receptacle parcels and accountables. The 3999 dataset that forms the basis for the updating process proposed in RM2017-8 also records separate delivery times for deviation parcels and accountables.

[45] Report on the City Carrier Street Time Study, Dkt. No. RM2015-7 (Dec. 11, 2014), at 43.

[46] *See* Notice and Order Establishing Docket Concerning City Carrier Special Purpose and Letter Route Costs and to Seek Public Comment, Dkt. No. PI2017-1 (May 31, 2017), at 3.

[47] *See* Responses of the United States Postal Service to Questions 1-7 of Chairman's Information Request No. 1, Dkt. No. PI2017-1 (Jun. 30, 2017), at 15 (response to Question 5).

The Postal Service appears to recognize that an update or revision of the established methodology is warranted.[48] UPS agrees that the current costing for special purpose routes is out-of-date and believes a new approach is long overdue. UPS also believes, however, that more clarity regarding SPRs is necessary before more constructive comments can be provided on this important issue.

In particular, the line between "special purpose" and "letter route" activities remains unclear.[49] Without clear definitions for these two sets of activities, the Commission and other interested parties will not be able to evaluate whether the Postal Service is accurately attributing its costs to products.[50] The treatment of "second run" delivery operations, for example, is especially unclear. The Postal Service has stated that sometimes carriers make "second runs" along their routes when too many parcels were present during the first run.[51] These "second run" routes appear to be wholly incremental to parcels, and thus costs associated with such routes should likely be attributed to parcels in their entirety.[52] The Postal Service, however, treats these routes

---

[48] *See* Responses of the United States Postal Service to Questions 1-7 of Chairman's Information Request No. 1, Dkt. No. PI2017-1 (Jun. 30, 2017), at 15 (response to Question 5: "This is not to suggest that an update or revision of the established methodology would not provide a more accurate representation of current activity proportions.").

[49] *See* United Parcel Service, Inc.'s Comments Regarding Proposal Five, Dkt. No. RM2017-9 (August 16, 2017), at 3.

[50] *See id*. at 4.

[51] Responses of the United States Postal Service to Questions 1-15, 19-20, and 23 of Chairman's Information Request No. 1, Dkt. No. RM2017-9 (Aug. 9, 2017), at 26 (Response to Question 15(b)).

[52] *See* Order No. 3506 Regarding UPS Proposals, Dkt. No. RM2016-2 (Sept. 9, 2016)), at 125 ("The Commission directs the Postal Service to use incremental costs as the basis for class-level and product-level attributable costs").

as if they were normal letter routes and applies the standard cost methodology to them.[53]

As these overflow routes increase in number, particularly during the Christmas peak season, the error introduced by treating them as standard letter mail routes becomes ever larger. Even though these routes may in reality be entirely caused by parcel volumes, a naïve application of the standard methodology would attribute only 3.7% of the costs associated with such routes to competitive products.[54]

UPS submitted a list of proposed questions related to Special Purpose Routes as "Appendix A" to its Comments on Proposal Five.[55] UPS believes that answers to these questions would help all Postal stakeholders to have a better understanding of the cost methodology associated with Special Purpose Routes.

CONCLUSION

UPS appreciates the work the Postal Service has done to investigate the single equation model. The preliminary results of its work confirm UPS's concern that Proposal Thirteen understates costs associated with parcel delivery. UPS aims to be constructive with suggestions on how to develop a single equation model that improves on Proposal Thirteen. UPS's concerns were raised two years ago, and notwithstanding the Postal Service's current objections to the single equation model, UPS believes that further development of the model is warranted.

---

[53] Responses of the United States Postal Service to Questions 1-15, 19-20, and 23 of Chairman's Information Request No. 1, Dkt. No. RM2017-9 (Aug. 9, 2017), at 26 (Response to Question 15(b)).

[54] *See* USPS-FY16-32. The Letter Route Distribution Key in the "Outputs to CRA" tab in CS06&7-Public-FY16.xlsx indicates that 3.7% of total letter route costs are attributed to competitive products.

[55] *See* United Parcel Service, Inc.'s Comments Regarding Proposal Five, Dkt. No. RM2017-9 (August 16, 2017), Appendix A.

The Commission should not accept continued delays in developing a single equation model, especially given that the mail mix is rapidly changing. Indeed, any approach that ameliorates the deficiencies in Proposal Thirteen should be carefully considered. For the foregoing reasons, UPS respectfully requests that the Commission:

1) Order the Postal Service to investigate the following alterations to its prototype "single equation" models:

   a. Extension of the dataset to encompass a broader temporal cross-section of the year and to a broader set of ZIP code days;

   b. Evaluate and test the four alternative "collections mail" approaches outlined in these comments: (1) Linear Feet (2) Reprogramming (3) Form 3999 and (4) CCCS;

   c. Re-evaluate the accountables variabilities as estimated following the work described in part (a), and consider the removal of "accountables" volume as a cost driver from the single-equation model, with appropriate adjustments.

   d. Consider opening a new docket or technical conference to vet UPS's proposals.

2) Issue Chairman's Information Requests Regarding Special Purpose Routes to develop themes proposed by UPS in this docket and in RM2017-9 (Proposal Five).

Respectfully submitted,

UNITED PARCEL SERVICE, INC.,

By /s/ Steig D. Olson

Steig D. Olson
Christopher M. Seck
David D. LeRay
Andrew Sutton
Quinn Emanuel Urquhart &
Sullivan, LLP
51 Madison Ave., 22nd Floor
New York, NY 10010
(212) 849-7152
steigolson@quinnemanuel.com
christopherseck@quinnemanuel.com
davidleray@quinnemanuel.com
andrewsutton@quinnemanuel.com

Attorneys for UPS

APPENDIX A

This Appendix presents the results of analytical work performed in support of the comments presented above, particularly with respect to the issue of multicollinearity and how best to address it.

Adding additional ZIP codes

A simple exercise that uses the diagnostic measures for multicollinearity relied upon by the Postal Service in its August 18 report generally supports the conclusion that adding ZIP codes may be more helpful than adding additional days for the same ZIP codes. While UPS only has access to data from 300 ZIP codes, and thus cannot be certain as to how the Condition Index (CI) and Variance Inflation Factor (VIF) would change if the dataset were expanded, the data that has been provided can be used to investigate how the CI and VIF change if the dataset is restricted. This exercise suggests the kinds of new data that might be most useful in mitigating concerns about multicollinearity.

The table below presents the Condition Index for the matrix of explanatory variables using a variety of different datasets that differ in terms of the number of days and the number of ZIP codes included. In all cases, the original single-equation model (with split parcel variables and no differentiation between FSS and non-FSS ZIP codes) was estimated. In general, for a given set of ZIP codes, the table indicates that multicollinearity is, in most cases, becoming gradually less severe as additional days are added to the dataset (moving from left to right in a given row of the table). However,

for a given set of days, the addition of ZIP codes (moving from top to bottom in a given column of the table) tends to reduce more significantly the degree of multicollinearity.[56]

Table 3: Condition Indices with Restrictions on Zip Codes and Observation Dates in Prototype Regression Model

Number of Zip Codes Included	July 1 -15	July	July + September
75	145.2	146.7	146.7
150	131.4	129.8	124.7
300	113.0	110.6	108.2

Sources: census_study_zips, study_dois_pa_vol_july, study_dois_pa_vol_july_sept.
Note: Zip code subsets are chosen based on the order of the zip codes in the dataset provided by USPS (for example, condition indices in the first row are calculated for a regression of the first 75 zip codes that appear in the dataset). To ensure that this is not driven by the order of the zip codes in the dataset, we have randomized the order 100 times and repeated the condition index calculation for each subset of the data. A similar table presents the average condition indices within each subset across these 100 random orderings and is qualitatively similar to this table. That table is provided in library reference UPS-PI2017-1/1.

This exercise suggests that data from additional ZIP codes would mitigate any multicollinearity concerns more effectively than the addition of September 2016 data has done.

Adding additional monthly data

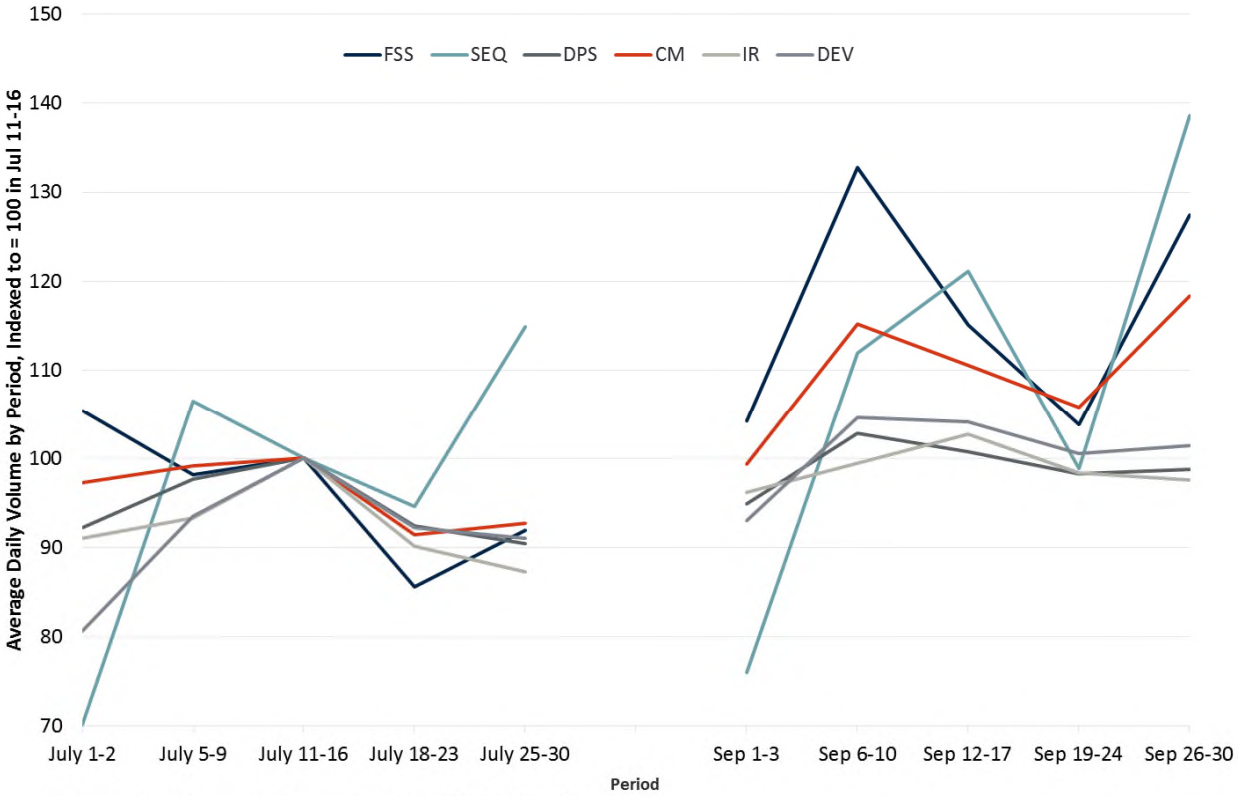
In its report, the Postal Service notes that "because the additional volume data contains the same patterns as the original volume data, addition of data will likely have only a minimal impact on the multicollinearity." This conclusion is based on the effect of adding September 2016 data to the original July 2016 data.[57]

[56] Similar tables, relying instead on the VIF diagnostic measure for multicollinearity, are consistent with the results presented here and are provided in library reference UPS-PI-2017-1/1.

[57] USPS Report at 14.

There are, however, indications that September adds little relevant variation to the original July dataset. Figure 1, below, plots the evolution of volume of the various major mail shapes over the course of the ten calendar weeks that are either partially or fully included in the expanded dataset that includes both July and September. For each of the major mail shapes included in the single-equation prototype, the graph depicts the daily average volume by week, using the 300 ZIP codes included in the dataset; those volumes have been indexed to 100 in the first full week of July.

Figure 1: Indexed Average Volume by Period and Mail Product



Source: study_dois_pa_vol_july_sept, from USPS-PI2017-1/2.
Notes: Three observations with negative deviation parcels volume are excluded.

The graph indicates that average volumes of several of the shapes ebb and flow in very similar ways over the period covered. Notably, DPS, in-receptacle parcels, and deviation parcels all tend to increase or decrease together from one week to the next.

FSS and CM roughly follow a similar, but more pronounced pattern to that of the first three volumes mentioned. Only SEQ mail seems to follow a pattern that differs in a meaningful way. It is also notable from this graph that the similar movements in several of these mail shapes continue into September. This is consistent with the Postal Service's point that the added volume data contain the same patterns as the original volume data, and is illustrative of why the addition of that particular data had only minimal impacts on the multicollinearity diagnostics.

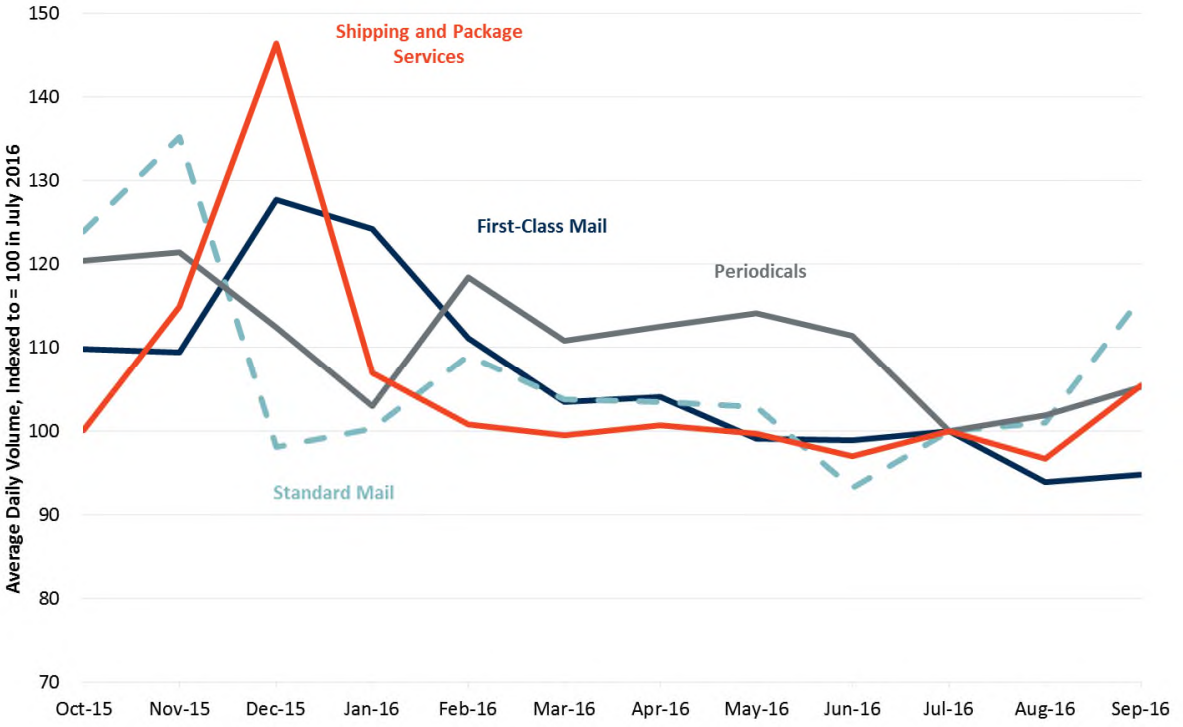
While UPS does not have access to data to see whether this pattern holds for other times of the year at the shape level, other public sources of data indicate that, at least at the product group level, other times of year may contain significantly more variation.

Figure 2 presents a similar graph, based on monthly volume estimates that are provided by the Postal Service in regular reports to the Commission.[58] The average daily volumes for these four product groups have been indexed to 100 for the month of July 2016; both figures use the same scale.[59]

[58] The data presented in these reports reflect estimates of system-wide (and not just city carrier) volume. Nevertheless, it is likely that the monthly variation in system-wide volumes is also reflected in delivery volumes on city carrier routes.

[59] While there are not generally one-to-one relationships between product groups and the mail shapes that are included as drivers in the prototype top-down equation, there are notable differences in the extent to which various mail shapes are represented in those product groups. First Class Mail and Standard Mail are both largely comprised of DPS and to a much lesser extent Cased Mail. Periodicals are heavily weighted towards Cased Mail, and Shipping and Package Services are largely comprised of deviation and in-receptacle parcels. All SEQ mail is Standard Mail, while FSS is heavily weighted towards both Standard Mail and Periodicals. See USPS-FY16-32, tab "Input DK" of CS06&7-Public-FY16.xlsx.

Figure 2: Indexed Average Daily Volume by Month and Product Group, FY16



Source: USPS Preliminary Financial Information (Unaudited) Reports, October 2015 - September 2016.
Notes: Average daily volumes equal monthly volume totals, divided by delivery days in the month, then indexed to 100 in July 2016.

One key implication of Figure 2 is that the changes from July to September are modest relative to those that occur during other parts of the year. It is also clear from Figure 2 that other periods of the year—notably October through February—are periods of time where the mail mix changes significantly, both with respect to July and to other months. The volumes of the various product groups often move in divergent ways from month to month. For example, Shipping and Package Service volumes peak in December, when Standard Mail volumes are close to their lowest. First-class mail volumes are near their peak in January, when Periodical volumes are near their lowest.

To the extent that the changes in product group mix (*i.e.*, the evolution over the course of the year in the volumes of Standard Mail, First Class Mail, Shipping and

Package Services, etc.) are at least partially reflected in changes to the mix of shapes (i.e. the evolution over the course of the year in the volumes of DPS, Cased Mail, in-receptacle parcels, etc.), the inclusion of other periods—beyond July and September—is likely to add more variation than did the addition of September, and thus holds promise for more reliable identification of marginal times and variabilities. In the past, the Postal Service has justified concentrating its data collection efforts in times of year that are typical or representative, and/or least likely to be disruptive.[60] However, the existence of PTR volume and DOIS street time data that are collected year round suggest this variation over the course of the year can now be exploited without disrupting regular operations. The potential for this variation to improve the reliability of variability estimates and minimize any multicollinearity concerns should be further investigated.

Discussion

As discussed above, the use of a dataset that covers a broader range of ZIP codes and times of the year may be effective in assuaging the Postal Service's concerns about multicollinearity. However, even if it is only partially effective, it is worth noting that multicollinearity does not render a model useless. Professor Ed Leamer has written that "[t]here is no pair of words that is more misused both in econometrics texts and in the applied literature than the pair 'multi-collinearity problem'."[61] Professor

[60] Responses of the United States Postal Service to Questions 1-11 of Chairman's Information Request No. 1, Dkt No. RM2017-8 (Aug. 2, 2017), at 11 ("To minimize disruptions, the district only initiates a set of route evaluations in a ZIP Code if they can be completed in a timely and efficient manner.")

[61] Edward E. Leamer, "Model Choice and Specification Analysis," in Zvi Griliches and Michael D. Intriligator, eds., Handbook of Econometrics, vol. I, North Holland Publishing Company, Amsterdam, 1983, at 300.

Christopher Achen elaborates: "[b]eginning students of methodology occasionally worry that their independent variables are correlated—the so-called multicollinearity problem. But multicollinearity violates no regression assumptions. Unbiased, consistent estimates will occur; and their standard errors will be correctly estimated. The only effect of multicollinearity is to make it hard to get coefficient estimates with small standard error."[62]

Particularly in the present context, the precision of individual coefficients is ultimately much less important than the precision of the resulting variabilities that are then used in the costing model. The variabilities are estimable functions of the estimated regression coefficients, as the Postal Service has correctly noted in the past.[63] In discussing multicollinearity concerns, econometricians frequently advocate a "do nothing" approach if the relevant *combination* of regression coefficients is statistically significant.[64] Thus, the Postal Service may not even be asking the right question in assessing the extent to which multicollinearity may or may not be a problem.

Furthermore, the criteria used by the Postal Service to evaluate the prototype model indicate that the established model for regular delivery time also displays the characteristics of multicollinearity. The Postal Service has noted that when assessing multicollinearity, variance inflation factors greater than ten may be problematic.[65] Similarly, the Postal Service report has cited a study by Belsley et al that concluded that

---

[62] CHRISTOPHER H. ACHEN, INTERPRETING AND USING REGRESSION 82-83 (1982).

[63] *See, e.g*. Report on the City Carrier Street Time Study, Dkt. No. RM2015-7 (Dec. 11, 2014), at 67.

[64] *See, e.g*., PETER KENNEDY, A GUIDE TO ECONOMETRICS 196 (6TH ED.) ; DAMODAR N. GUJARATI, BASIC ECONOMETRICS 363-64 (2004).

[65] USPS Report at 20.

"a Condition Index greater than 30 indicates moderate dependencies among the right-hand-side variables, and a value approaching 100 indicates strong dependencies."[66] However, by these metrics, the regressions presented by the Postal Service in the Report on the City Carrier Street Time Study in Docket No. RM2015-7 also suffered from similarly high levels of multicollinearity.

Revisiting the "Initial Estimation of the Regular Delivery Time Equation" (p. 65 of the Report on the City Carrier Street Time Study) with these same multicollinearity diagnostics indicates that the Condition Index for that regression is 84.6, and that 21 of 33 (or 64%) of the coefficients have variance inflation factors above ten. The Postal Service then removed all terms that were not statistically significant in their initial regression, leading to a slight improvement in the multicollinearity diagnostics when applied to the final model. Nevertheless, the Condition Index for the final regression is 74.4, and 17 of 26 (or 65%) of the coefficients have variance inflation factors above ten.[67]

In short, UPS believes that the possibility to exploit the increased variation in a dataset based on operational data from a greater number of ZIP codes and a more varied time period holds enormous promise for assuaging the Postal Service's concerns with respect to multicollinearity. In any case, the Commission should keep in mind that

---

[66] *Id*. at 21.

[67] While UPS' external consultants have successfully replicated the variance inflation factors reported in Table 27 on p. 71 of the Report on the City Carrier Street Time Study, they have not been able to replicate the Condition Indices reported on pp. 73 and 75 of that Report, and the library reference in Docket RM2015-7 does not appear to include the SAS program that produced those values. The diagnostic values reported here were produced by applying the SAS commands used by the Postal Service to diagnose multicollinearity in USPS-PI2017-1/2 to the regular delivery time regressions estimated in USPS-RM2015-7/1. The programs and results used by UPS' external consultants are provided in UPS-PI2017-1/1.

the regression models upon which the existing methodology relies are far from immune from those same concerns, and that leading econometricians and statisticians have argued that concerns about multicollinearity like those the Postal Service have expressed are often misplaced, overstated, or both.